Hydrogen bonding part 46: A review of the correlation and prediction of transport properties by an LFER method: physicochemical properties, brain penetration and skin permeability†

Michael H Abraham,[1]* Harpreet S Chadha,[1] Filomena Martins,[2] Robert C Mitchell,[3] Michael W Bradbury[4] and Julie A Gratton[4]

[1] *Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK*

[2] *Grupo de Estrutura e Reactividade Quimica, Department of Chemistry, University of Lisbon, Cc. Bento da Rocha Cabral 14, 1250 Lisboa, Portugal*

[3] *SB Pharmaceuticals, New Frontiers Science Park (North), 3rd Avenue, Harlow, Essex, UK*

[4] *Physiology Group, Biomedical Sciences Division, King's College London, Strand, London WC2R 2LS, UK*

Abstract: A number of solute descriptors that relate to the ability of a solute to take part in solute–solvent interactions have been identified, quantified and incorporated into a multiple linear regression equation. This general solvation equation can then be used for the correlation and prediction of solute effects in transport processes, that is, processes in which the main step is either the equilibrium transfer, or the rate of transfer, of a solute from one phase to another. Examples discussed include the solubility of gases and vapours in water, various water-solvent partitions, blood-brain distribution, brain perfusion, and skin permeability.

Keywords: LFER; QSAR; brain penetration; skin permeation; solvation; hydrogen bonding

1 INTRODUCTION

A large number of transport-related processes are of importance in pharmaceutical and environmental chemistry. These processes involve either the equilibrium transfer or the rate of transfer of a compound from one phase to another phase. As regards pharmaceutical or biochemical processes, we deal with passive transfer only. Equilibrium transfer is controlled by the standard Gibbs energy of the compound in the two phases, which in turn is related to the Gibbs energy of solvation of the gaseous compound in the two phases, as shown in Fig 1. Our general method of approach is based on the assumption that certain properties or descriptors of a given compound will be of importance in solvation of the compound, not just in one particular solvent phase, but in solvent phases in general. Thus if the polarisability, for example, of a compound is an important factor in solvation in solvent ethanol, we expect it to be relevant as regards solvation in chloroform, or benzene, or other solvents. Naturally, a compound property such as acidity will influence solubility in basic solvents, but will be redundant in non-basic solvents. However, with such caveats in mind, it was our intention to attempt to define and obtain com-

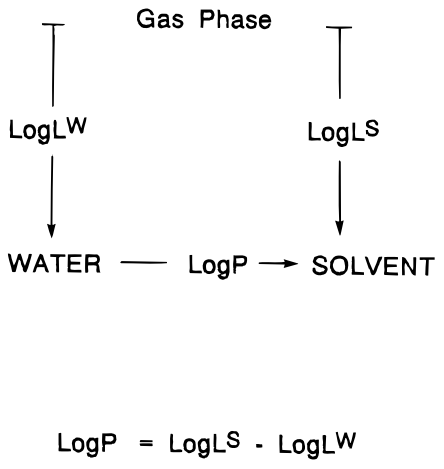

Figure 1. The relationship of gaseous solubility in water and a solvent to partition between water and the solvent.

* Correspondence to: MH Abraham, Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ.

† Based on a paper presented at the meeting 'Uptake of Agrochemicals and Pharmaceuticals', organised by AH Cobb, AM Stalker and DJ Livingstone on behalf of the Pesticides Group and the Physicochemical and Biophysical Panel of the SCI and the Strathclyde Institute for Drug Research and held on 9 December 1997 at 14/15 Belgrave Square, London

(*Received 19 January 1998; revised version received 19 June 1998; accepted 29 June 1998*)

pound descriptors that would, in different combinations, model solvation in various solvent phases. It is then clear, from Fig 1, that such descriptors would also enable transfer from one phase to another to be modelled. A similar argument can be constructed for rates of transfer, but now in terms of initial state and transition state solvation. Since we shall always consider rather dilute solutions of compounds, we shall refer to 'solutes' rather than to compounds.

The original work of Kamlet and Taft, and coworkers[1,2] had shown that it was indeed possible to define a rather small set of descriptors that could be combined in a linear way for the correlation of solute properties. After considerable preliminary work,[3,4] we eventually succeeded in constructing a new and more rigorous set of five solute descriptors,[5–9] specified as follows. R_2 is an excess molar refraction that is obtained from refractive index for solutes that are liquid at 20°C. For solids, the refractive index of the hypothetical liquid at 20°C can be calculated, or R_2 can be obtained by the summation of fragments or substructures. π_2^H is the dipolarity/polarisability that can be obtained from gas-liquid chromatographic measurements on polar stationary phases, or more generally from water–solvent partition coefficients. $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ are the overall or effective hydrogen-bond acidity and basicity that are most easily obtained from water–solvent partitions, and V_x is the McGowan characteristic volume[10] that can easily be calculated from bond and atom contributions;[8] in this work V_x is in units of (mol dm^{-3})/100. It should be noted that there is a very clear distinction between the $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ descriptors, and measures of proton acidity or basicity. In Table 1 are given some values of the hydrogen-bond descriptors $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$, together with values of pK_a and pK_{Hb} for the solutes in water. Across different families of compound, there is little connection between the two sets of data.

The range of solutes for which descriptors are currently available is now quite large, and encompasses compounds as far apart as helium, hydrogen, nitrogen, etc. on the one hand, and drugs, environmental pollutants and pesticides such as barbiturates, zolantidine, steroids, PCBs, DDT and dieldrin on the other hand. In Table 2 is a list of the number of compounds for which our descriptors have been established.

These solute descriptors can be combined into a linear free energy relationship, equation (1) where the dependent variable, log SP, is a set of solute properties in a given system, and the independent variables are the five solute descriptors. The coefficients in eqn (1) are found by the method of multiple linear regression analysis.

$$\log SP = c + r \cdot R_2 + s \cdot \pi_2^H + a \cdot \Sigma\alpha_2^H + b \cdot \Sigma\beta_2^H + v \cdot V_x \quad (1)$$

Table 1. Comparison of hydrogen-bond acidity and basicity with proton acidity and basicity

Solute	$\Sigma\alpha_2^H$	pK_a
Thiophenol	0.09	6.5
Ethanol	0.37	15.9
Phenol	0.60	10.0
Acetic acid	0.61	4.8
Solute	$\Sigma\beta_2^H$	pK_{Hb}
Acetonitrile	0.32	−10.1
4-Nitroaniline	0.38	1.0
Pyridine	0.52	5.2
Acetamide	0.68	0.0
Triethylamine	0.79	10.7
Dimethylsulfoxide	0.88	−1.8

Since the descriptors in equation (1) have been constructed so as to correspond to particular solute–solvent interactions, the coefficients will encode the interaction properties of the solvent phase; if transfer between two phases is considered, then the coefficients will indicate the difference in interaction properties of the two phases. The r-coefficient refers to the tendency of the solvent phase to interact through π- and σ-electron pairs, the s-coefficient to the phase polarisability/dipolarity, the a-coefficient to the phase hydrogen-bond basicity (because acidic solutes will interact with basic phases) and the b-coefficient to the phase hydrogen-bond acidity. The v-coefficient will relate to the phase hydrophobicity, which itself can be broken down into a cavity term and a general dispersion interaction term.

The general equation (1) can be used in several distinct ways. Firstly, it can be applied to the correlation and then to the prediction of some particular solute property. Secondly, the coefficients in equation (1) can form the basis of a general method of characterising both physicochemical and biochemical processes. Thirdly, a given regression equation can be analysed term-by-term in order to isolate and to quantify the particular interactions that influence the process under consideration. It is the aim of this paper to set out the results of such analyses for a number of important processes.

Table 2. Solute descriptors used in the general equation (1)

Code	*Descriptor*	*No*[a]	*Max*[b]	*Min*[b]
R_2	Excess molar refraction	3410	4.62	−1.37
π_2^H	Dipolarity/polarisability	2880	5.60	−0.54
$\Sigma\alpha_2^H$	Hydrogen-bond acidity	3730	2.10	0.00
$\Sigma\beta_2^H$	Hydrogen-bond basicity	2560	4.52	0.00
V_x	McGowan volume	>3600[c]	8.56	0.07

[a] The number of solutes for which the descriptor is currently available.
[b] The maximum and minimum values of the given descriptor.
[c] The volume, V_x, can easily be calculated for any structure.

2 APPLICATIONS

2.1 Physicochemical processes

When equation (1) is applied to a set of physicochemical data that are free-energy-based, we can refer to it as a linear free energy relationship (LFER). As an example of the solvation properties of a single phase, we can consider gas–water partition coefficients, or Ostwald solubility coefficients, of solutes defined through equation (2):

$$L = [\text{conc. in solvent}]/[\text{conc. in the gas phase}] \quad (2)$$

If the same units of concentration are used in water and the gas phase, eg mol dm^{-1}, L is a dimensionless quantity. With due regard to units, L is the inverse of Henry's constant. Abraham *et al*[11] applied equation (1) to a set of 408 log L values in water at 298 K, denoted as log L^W, and obtained the LFER equation (3):

$$\begin{aligned} \log L^W = &-0.994 + 0.577R_2 + 2.549\pi_2^H \\ &+ 3.813\Sigma\alpha_2^H + 4.841\Sigma\beta_2^H \\ &- 0.869V_x \qquad (3) \end{aligned}$$

$$n = 408,\ r^2 = 0.9952,\ sd = 0.151,\ F = 16\,810$$

Here, and elsewhere, n is the number of data points, r is the correlation coefficient, sd is the standard deviation in the dependent variable, and F is the Fisher F-statistic.

It is important to note that the coefficients of the first four descriptors are all positive, as required by physicochemical considerations. Since there are no solute–gas phase interactions, any solute–water interactions must lead to an increase in solubility and hence to a positive coefficient. The v-coefficient will be the resultant of an unfavourable cavity effect in which water–water interactions are broken, and a favourable general dispersion solute–water interaction. In the case of solvent water, the unfavourable cavity effect dominates, thus leading to a negative v-coefficient. For all non-aqueous solvents we have studied, the v-coefficient is always positive. Water is unique in that large gaseous solutes dissolve to a lesser extent than small solutes, whereas in non-aqueous solvents it is the larger solutes that dissolve to the greater extent.

Once equation (3) is constructed, it can obviously be used to predict log L^W, or Henry's constant, for any solute for which the descriptors are available. But also, the coefficients in equation (3) now characterise the solvent. In particular the three s-, a- and b-coefficients are all very large, so that we can conclude that bulk water is dipolar, a good hydrogen-bond base ($a = 3.81$) and a good hydrogen-bond acid ($b = 4.84$). A term-by-term analysis of equation (3) can be carried out to show exactly how solute properties influence the gaseous solubility. Descriptors for a number of typical solutes are in Table 3, and results are in Table 4. The first two entries show the comparatively small effect of increase in solute volume. The remaining entries are for solutes of about the same volume, so that effects on log L^W are due to the other four terms. The main influences involve the three descriptors π_2^H, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$. Solutes that are dipolar, such as pentan-2-one, or are polarisable, such as toluene, will give rise to $s \cdot \pi_2^H$ terms that can be as large as two log units. Even more important are the acidity and basicity terms: the $a \cdot \Sigma\alpha_2^H$ term can be as large as three log units for the strong hydrogen-bond acid, hexafluoropropan-2-ol (HFIP, $(CF_3)_2CHOH$), and strong hydrogen-bond bases will be very soluble through the $b \cdot \Sigma\beta_2^H$ term, as for butylamine or 4-acetylpyridine. There are some surprising effects, however, mostly due to the very large b-coefficient. Thus the basicity of propanoic acid is almost as important as the acidity in producing the very large log L^W, and for butan-2-ol the solute basicity term contributes much more than the solute acidity.

We have dealt with gaseous solubility in water in some detail, because only one solvent phase is involved. For partition between phases, the coefficients in equation (1) will represent the difference in properties in the two phases. As an example, we use the well-known partition between water and octanol defined through equation (4):

$$P_{oct} = [\text{conc. in octanol}]/[\text{conc. in water}] \quad (4)$$

In the case of a solute that can (partially) ionise in

Solute	R_2	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$	V_x
Ethane	0.00	0.00	0.00	0.00	0.390
Pentane	0.00	0.00	0.00	0.00	0.813
Pentan-2-one	0.143	0.68	0.00	0.51	0.829
n-Butylamine	0.224	0.35	0.16	0.61	0.772
Propanoic acid	0.233	0.65	0.60	0.45	0.606
Butan-2-ol	0.217	0.36	0.33	0.56	0.731
Hexafluoropropan-2-ol	−0.240	0.55	0.77	0.10	0.696
Toluene	0.601	0.52	0.00	0.14	0.857
Phenol	0.805	0.89	0.60	0.30	0.775
4-Acetylpyridine	0.771	1.13	0.00	0.84	0.973

Table 3. Descriptors for some solutes

Solute	$r \cdot R_2$	$s \cdot \pi_2^H$	$a \cdot \Sigma\alpha_2^H$	$b \cdot \Sigma\beta_2^H$	$v \cdot V_x$	Log L^W	
						Calc	Obs
Ethane	0.00	0.00	0.00	0.00	−0.34	−1.33	−1.34
Pentane	0.00	0.00	0.00	0.00	−0.71	−1.70	−1.70
Pentan-2-one	0.08	1.73	0.00	2.47	−0.72	2.57	2.50
n-Butylamine	0.14	0.89	0.61	2.96	−0.67	2.94	3.11
Propanoic acid	0.13	1.66	2.29	2.18	−0.53	4.74	4.74
Butan-2-ol	0.12	0.92	1.26	2.72	−0.63	3.39	3.38
Hexafluoropropan-2-ol	−0.14	1.40	2.94	0.48	−0.60	3.09	2.76
Toluene	0.35	1.33	0.00	0.68	−0.74	0.63	0.65
Phenol	0.46	2.27	2.29	1.45	−0.67	4.81	4.85
4-Acetylpyridine	0.44	2.88	0.00	4.07	−0.84	5.56	5.59

Table 4. Factors influencing gaseous solubility in water at 298 K

water, such as carboxylic acids or strong proton-bases, P_{oct} refers to the partition of the neutral species. The water–octanol system is particularly important, because it is widely used as a model system for a variety of biochemical and environmental processes. Values of $\log P_{oct}$ can be correlated[12] through the LFER equation:

$$\log P_{oct} = 0.088 + 0.562R_2 - 1.054\pi_2^H + 0.034\Sigma\alpha_2^H - 3.460\Sigma\beta_2^H + 3.814V_x \quad (5)$$
$$n = 613,\ r^2 = 0.9948,\ sd = 0.116,\ F = 23\,162$$

The coefficients in equation (5) now refer to the difference in properties of octanol, or more correctly wet octanol, and water. The negative *s*- and *b*-coefficients arise because water is more dipolar and is a stronger hydrogen-bond acid than wet octanol. The almost-zero *a*-coefficient shows that the hydrogen-bond acidity of solutes plays no part in the partition, and also that water and wet octanol have the same hydrogen-bond basicity. Finally, the large *v*-coefficient indicates that octanol is a very hydrophobic phase. For nonpolar solvents, the *a*-coefficient (like the *b*-coefficient) is large and negative; for example, partition coefficients between water and hexadecane, P_{16}, are given[12] by:

$$\log P_{16} = 0.087 + 0.667R_2 - 1.617\pi_2^H - 3.587\Sigma\alpha_2^H - 4.869\Sigma\beta_2^H + 4.433V_x \quad (6)$$
$$n = 370,\ r^2 = 0.9964,\ sd = 0.124,\ F = 20\,236$$

Since the *a*-coefficient is related to the difference in hydrogen-bond basicity between hexadecane and water, it is no surprise that it is large and negative. Indeed, since hexadecane has no acidity or basicity at all, we expect that the α- and *b*-coefficients in equation (6) will be the most negative that can be observed. Within a reasonable experimental error, this is the case.[9,12] Another process that has aroused attention, is the cyclohexane-to-octanol partition, defined by Seiler[13] as:

$$\Delta \log P_{cyc} = \log P_{oct} - \log P_{cyc} \quad (7)$$

Actually, it makes little difference if alkanes are used instead of cyclohexane. For the case where hexadecane[12] is used:

$$\Delta \log P_{16} = -0.072 - 0.093R_2 + 0.528\pi_2^H + 3.655\Sigma\alpha_2^H + 1.396\Sigma\beta_2^H - 0.521V_x \quad (8)$$
$$n = 288,\ r^2 = 0.9669,\ sd = 0.173,\ F = 1646$$

The main factor that influences $\Delta \log P_{16}$ is solute hydrogen-bond acidity, but other factors are still important.

Equations such as equation (5) and equation (6) have been constructed for numerous water–solvent partitions, so that once descriptors have been assigned to a given solute it is trivial to predict a large number of partition coefficients.[9,12]

The general equation (1) has been applied to several other physicochemical processes, including HPLC, TLC and other chromatographic methods, solid-phase extraction and partitioning into micelles. At the moment we are investigating the aqueous solubility of liquids and solids, and water–soil partitioning, and there seems little doubt that very many physicochemical transport processes can usefully be studied via equation (1).

2.2 Biochemical processes

A number of quantitative structure–activity relationships (QSARs) have been constructed through equation (1). One of the first biochemical processes we examined was narcosis of the tadpole by aqueous solutes.[14] Overton,[15] in 1901, had reported aqueous concentrations, C_{nar}, that just brought about narcosis, and Lipnick[16] had made these data readily available. Application of equation (1) yielded the QSAR:

$$\log(1/C_{nar}) = 0.579 + 0.825R_2 - 0.334\pi_2^H - 2.871\Sigma\beta_2^0 + 3.097V_x \quad (9)$$
$$n = 84,\ r^2 = 0.9467,\ sd = 0.246,\ F = 351$$

The $\Sigma\beta_2^0$ descriptor in equation (9) takes into account the variable basicity of a few solutes.[6] If values of C_{nar} from several other sources were also included, the resulting equation was:

$$\log(1/C_{nar}) = 0.595 + 0.805R_2 - 0.725\pi_2^H - 2.489\Sigma\beta_2^0 + 3.341V_x \quad (10)$$
$$n = 114,\ r^2 = 0.9063,\ sd = 0.341,\ F = 263$$

Quite remarkably, only five compounds were excluded from equation (10), of which three were used at such high concentration that osmotic effects were likely to interfere. Hence tadpole narcosis is a very general phenomenon. The solute factors that influence narcosis are self-evident from equation (10): the two main factors are solute hydrogen-bond basicity that decreases toxicity (ie increases the narcotic concentration) and solute volume that increases toxicity.

2.3 Brain penetration

The blood–brain barrier (BBB) is formed by high-resistance tight junctions between the endothelial cells in the cerebral microvessel walls, and between epithelial cells of the choroid plexus. In order to pass into the brain, compounds must undergo passive diffusion via a transcellular route, or some form of active transport.[17]

There are a number of measures of what may loosely be described as 'brain penetration' or the 'ability of compounds to cross the BBB'. Some of these measures have been based on biological activity,[18] or on drug–receptor interactions.[19,20] Such measures must include various effects, and, although useful in some contexts, are not suitable for any analysis of factors that influence just passive transport across the BBB. Two particular well-defined measures of 'brain penetration' that are amenable to quantitative analysis are available, however. These are (i) the steady-state distribution of a compound between blood and brain, and (ii) the rate of permeation of a compound from blood or saline through the BBB. We consider these in turn.

2.3.1 *Blood-brain distribution*

The distribution of a compound between blood and brain can be described by an equilibrium constant denoted as BB:

$$BB = [\text{conc in brain}]/[\text{conc in blood}] \quad (11)$$

The main experimental determination of values of BB by an in-vivo method on rats was carried out by Young, Mitchell *et al*,[21] (YM) who obtained BB values for 30 drug-type compounds by a radioassay method. For a subset of 20 compounds, they showed that there was little correlation of log BB with log P_{oct} ($r^2 = 0.190$), a rather better correlation with the water–cyclohexane partition coefficient as log P_{cyc} ($r^2 = 0.536$), and a more reasonable correlation with the $\Delta\log P_{cyc}$ parameter ($r^2 = 0.691$).[21]

We used all 30 of the YM set together with 35 indirectly determined BB values[22] and applied equation (1) to the entire set. Of the 65 solutes, there were eight outliers, but the remaining 57 solutes yielded the correlation:[23]

$$\log BB = -0.038 + 0.198R_2 - 0.687\pi_2^H - 0.715\Sigma\alpha_2^H - 0.698\Sigma\beta_2^H + 0.995V_x \quad (12)$$
$$n = 57,\ r^2 = 0.9162,\ sd = 0.197,\ F = 99$$

The fact that there are eight outliers of the YM set to equation (12) is not surprising. The distribution experiments may take up to several hours before equilibrium is established, and any biological degradation of a compound would result in an incorrect BB value, as obtained from a radioassay.

Equation 12 shows, for the first time, exactly the solute factors that govern BB values. Increase in R_2 (weakly) and V_x (greatly) increases distribution into the brain, whereas increase in π_2^H, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ all decrease brain distribution. Equation 12 puts all this on a quantitative basis, so that it is now possible to design drugs with increased or decreased BB values on a rational basis. As an example, we show in Table 5 how a particular substitution on zolantidine, Fig 2, affects the solute descriptors and hence the expected values of blood–brain distribution.[24]

A comparison of equation (12) with equation (5) is instructive in that it is now apparent why log P_{oct} is a poor descriptor for log BB. In eqn (5) the term in $a \cdot \Sigma\alpha_2^H$ is not significant, whereas in the log BB equation this term is as large as any other term. Hence solutes that are hydrogen-bond acids will always be out of line when log BB is plotted against log P_{oct}. For 49 of the 57 solutes in equation (12), log P_{oct} values were available, and the following regression was constructed:[23]

$$\log BB = -0.376 + 0.216 \log P_{oct} \quad (13)$$
$$n = 49,\ r^2 = 0.2696,\ sd = 0.498,\ F = 17$$

Table 5. Calculated descriptors for substituted zolantidine,[a] and the predicted values of log *BB*

X	R_2	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$	V_x	log BB
H	2.69	2.64	0.40	1.38	2.9946	0.42
Et	2.69	2.67	0.40	1.39	3.2764	0.68
OMe	2.80	2.90	0.40	1.54	3.1942	0.35
OH	2.82	3.06	0.99	1.48	3.0533	−0.27
NH_2	3.01	3.11	0.65	1.64	3.0944	−0.10

[a] See Fig 2 for structure.

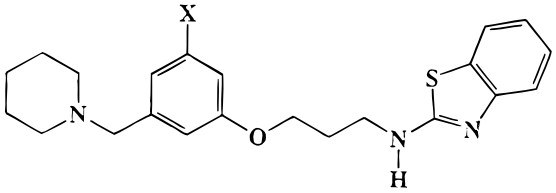

Figure 2. Substituted zolantidines. (Study reported in Table 5). (X = H, zolantidine)

A plot of log *BB* observed vs log *BB* calculated on equation (12) is shown in Fig 3, where there is random scatter about the line of identity. A similar plot on equation (13) is shown in Fig 4; it is clear that solutes fall onto two main lines, with solutes that are strong hydrogen-bond acids forming the lower line. This is exactly as predicted from equation (5) and equation (12), because hydrogen-bond acids will decrease the value of log *BB*, but will not affect log P_{oct}. It is therefore very clear that log P_{oct} will be a poor predictor of log *BB*. On the other hand, equation (12) has been shown to predict log *BB* values for a number of solutes not used to set up equation (12).[25]

There have been several other reported attempts to correlate log *BB* values with various descriptors. That of van de Waterbeemd and Kansy[26] has been shown[27] not to yield reliable predictions and the use of HPLC capacity factors on an artificial immobilised membrane produced a very poor correlation

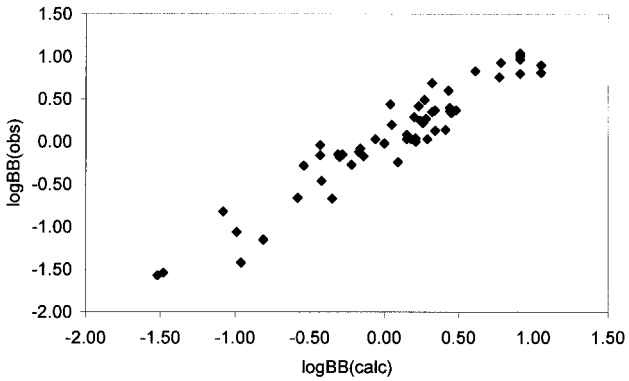

Figure 3. A plot of log *BB* observed vs log *BB* calculated on equation (12).

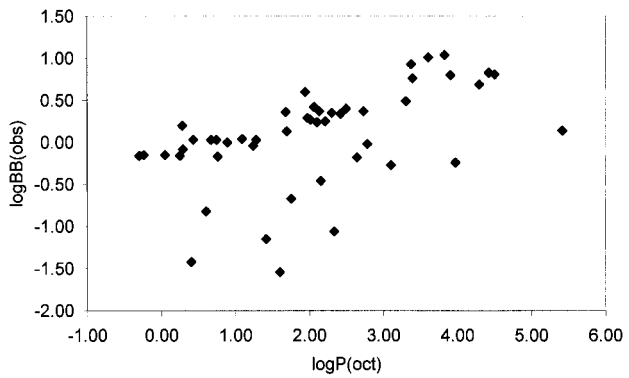

Figure 4. A plot of log *BB* observed vs log P_{oct}, see equation (13).

Table 6. Summary of the more general equations for the calculation of blood–brain distribution, as log *BB*[a]

Descriptor	*Eqn*	*n*	*sd*	*Ref*
log P_{oct}	13	49	0.50	23
Δ log *P*	–	32	0.27	23
ΔG^0_W	14	55	0.41	29
As equation (1)	12	57	0.20	23

[a] The range in observed log *BB* values is about 3 log units.

equation.[28] Lombardo *et al*[29] used the calculated Gibbs energy of solvation in water, ΔG^0_W in kcal mol^{-1}, of gaseous solutes as a descriptor and obtained the equation:

$$\log BB = 0.43 + 0.054\ \Delta G^0_W \qquad (14)$$
$$n = 55,\ r^2 = 0.67,\ sd = 0.41,\ F = 108$$

Log *BB* values for a number of compounds outside the training set of 55 solutes were quite well predicted through equation (14). A summary of the more important equations for log *BB* is in Table 6.

2.3.2 *Brain perfusion*

In a typical perfusion experiment, the solute in saline or in blood is perfused into the internal carotid artery and the rate of uptake into the brain is obtained by a radioassay method in which a series of rats are sacrificed at various intervals of time. Unlike the *BB* distribution experiments, the time scale in the perfusion technique is very short – no more than a few minutes – so that untoward processes such as biological degradation are less likely to take place.

Care must be taken in the comparison of one set of perfusions with another. The experimental results may depend on differences in technique and will certainly depend on the different perfusate solutions used. Thus perfusion from blood, or saline containing albumin, or pure (buffered) saline will not be the same.

In our work,[30] protein-free saline was used, and the permeation rates of 18 varied solutes were determined;[31] they are expressed as a permeability surface area product, *PS* in cm^3 s^{-1} g^{-1}. Since some of the compounds were partially ionised in saline, all the PS-products have been corrected for ionisation and refer to perfusion of the neutral species. Application of equation (1) to the 18 log *PS* products gave the correlation:

$$\log PS = -1.21 + 0.77R_2 - 1.87\pi^H_2 - 2.80\Sigma\beta^H_2 + 3.31V_x \qquad (15)$$
$$n = 18,\ r^2 = 0.9526,\ sd = 0.481,\ F = 65$$

This shows that increase in R_2 and V_x increases the permeation rate, and that increase in the solute

hydrogen-bond basicity decreases the rate. Unlike the equation for log BB, equation (12), solute hydrogen-bond acidity has no effect on log PS. Hence, as suggested above, different measures of 'brain penetration' will not be modelled in the same way.

Because equation (15) lacks any term in $a \cdot \Sigma\alpha_2^H$, it more resembles the log P_{oct} equation, and so we expect that log P_{oct} might be a suitable descriptor for log PS, even though it is a poor descriptor for log BB. Indeed, for the same 18 solutes we find:

$$\log PS = -2.28 + 0.69 \log P_{oct} \qquad (16)$$
$$n = 18,\ r^2 = 0.7770,\ sd = 0.94,\ F = 55$$

Plots of log PS observed vs log PS calculated on equation (15) and equation (16) are shown in Fig 5 and Fig 6. On equation (15), there is random scatter about the line of identity, but inspection of Fig 6 suggests that there might just be a parabolic relationship of log PS with log P_{oct}. A statistical analysis indicates otherwise.[31]

Timmermans *et al*[32] have suggested a parabolic relationship with log P_{oct} for permeation from saline, but other workers find only a linear relationship for permeation from plasma (corrected for protein binding),[33] or from saline containing bovine serum albumin.[34] In any case, it seems rather clear that the solute factors governing brain permeation are not quantitatively the same as those influencing blood–brain distribution.

2.4 Skin permeation

The permeation of solutes from aqueous solution through the human skin is of considerable environmental importance, and a number of algorithms have been described for the permeability coefficient, as log K_p. Many of these have been restricted to particular groups of compound, and so are not general enough for many practical purposes. There are a number of algorithms, however, that cover a wide range of compound type; all the given equations in log K_p are functions of log P_{oct} and solute molecular weight.[35–39] Wilschut *et al*[39] have compared various equations, using a data set of 99 solutes. Unfortunately, for an important group of steroids, they used log K_p values originally determined by Scheuplein *et al*.[40] Recent work indicates that these values are not correct.[41]

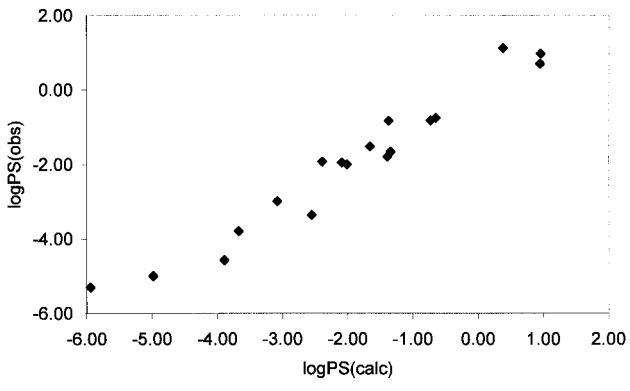

Figure 5. A plot of log *PS* observed vs log *PS* calculated on equation (15). Reproduced from the *Journal of Pharmacy and Pharmacology* by kind permission of the Editor.

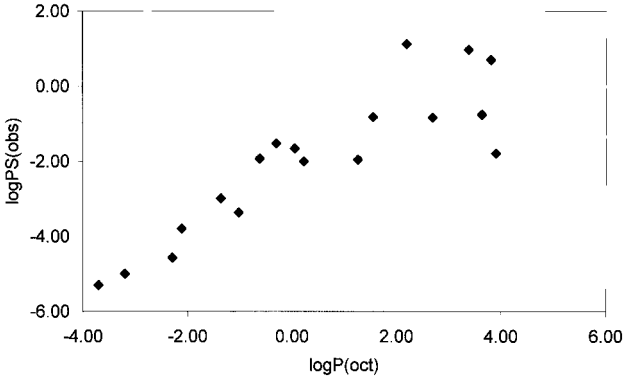

Figure 6. A plot of log *PS* observed vs log P_{oct}, see equation (16). Reproduced from the *Journal of Pharmacy and Pharmacology* by kind permission of the Editor.

Originally, we also used the Scheuplein value for steroids,[42] but subsequently analysed a large and varied data set that included both the Scheuplein data for steroids,[43] and the recent recommended values of log K_p by Johnson *et al*.[41] Permeability coefficients, K_p in cm s^{-1}, were assembled for the permeation of human skin by 47 aqueous solutes covering a wide range of compound type, but excluding any steroids. Application of equation (1) yielded the correlation:

$$\log K_p = -5.241 + 0.437R_2 - 0.410\pi_2^H - 1.631\Sigma\alpha_2^H - 3.286\Sigma\beta_2^H + 2.012V_x \qquad (17)$$
$$n = 47,\ r^2 = 0.9567,\ sd = 0.197,\ F = 181$$

Incorporation of the log K_p values suggested by Johnson *et al*[41] led to a correlation that was very close indeed:

$$\log K_p = -5.132 + 0.439R_2 - 0.489\pi_2^H - 1.478\Sigma\alpha_2^H - 3.442\Sigma\beta_2^H + 1.941V_x \qquad (18)$$
$$n = 53,\ r^2 = 0.9577,\ sd = 0.213,\ F = 213$$

Hence we can demonstrate that the Johnson values are compatible with log K_p values for non-steroid compounds. A plot of observed and predicted values of log K_p on equation (18) is shown in Fig 7; there is a reasonable straight line that includes the six steroids in the Johnson group. Previously[43] we considered five steroids studied by Johnson *et al*,[41] but now take an average value of log K_p for hydrocor-

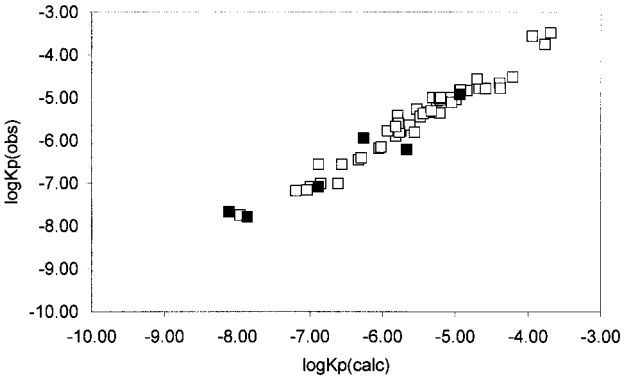

Figure 7. A plot of log K_p observed vs log K_p calculated on equation (18). (□) Nonsteroids, (■) six Johnson steroids. Reproduced from the *Journal of Pharmacy and Pharmacology* by kind permission of the Editor.

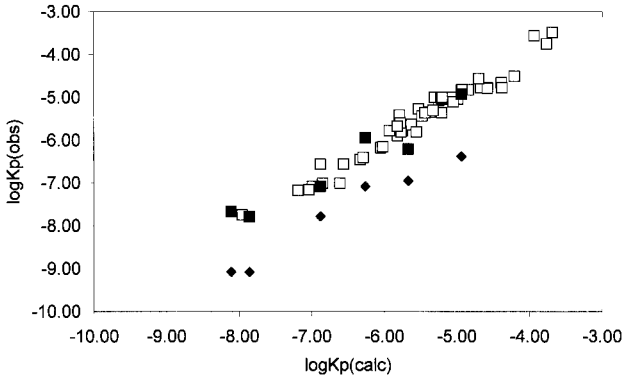

Figure 8. As for Fig 7; (◆) Scheuplein values of log K_p. Reproduced from the *Journal of Pharmacy and Pharmacology* by kind permission of the Editor.

tisone, − 7.67, to give a total of six steroids. In Fig 8 is a similar plot to that shown in Fig 7, but now including the Scheuplein log K_p values for the six steroids that Johnson studied. Note that an average value for hydrocortisone is shown in Figs 7 and 8. The log K_p values observed by Scheuplein *et al*[40] are all much lower than those found by Johnson *et al*.[41] It is not just these six common steroids that are out of line, but all the steroids studied by Scheuplein *et al*[40] have log K_p values lower than we calculate through equation (18), (see Table 7). Quite clearly, use of the Scheuplein data set would yield a different, and we think incorrect, regression equation.

We can interpret equation (18) in our usual way: the rate of permeation is increased by an increase in R_2 and especially in V_x, and the rate of permeation is decreased by an increase in π_2^H, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$. Since equation (18) is based on a wide variety of compound type, it should be useful in the prediction of log K_p values for compounds of environmental interest.

We did find,[43] however, that a set of steroid esters studied by Anderson *et al*,[44] did not conform to equation (18). These esters are all much larger than the solutes included in equation (18) (even the steroids) and it is possible that very large solutes permeate by a different mechanism (Johnson, ME, pers comm.)

3 DETERMINATION OF DESCRIPTORS

Any application of the general equation (1) depends on the availability of the solute descriptors. Although we have to hand descriptors for quite a large number and variety of compounds, (Table 2), there will always be a need to determine descriptors for further compounds. We have given an indication of the various methods that can be used,[9] as follows. Firstly, the descriptor V_x can be calculated quite simply for any structure from the molecular formula and the number of rings in the molecule, using the algorithm of Abraham for the number of bonds in a molecule.[8] Secondly, the R_2 descriptor can be calculated from the refractive index at 20°C, using either the observed refractive index for a liquid, or a calculated refractive index for a solid (this is the refractive index calculated for the hypothetical liquid). Since the R_2 descriptor is based on molar refraction, it can

Table 7. Comparison of observed values of log K_p for steroids, with values calculated on equation (16)

Steroid	*Observed log K_p*		*Calculated log K_p eqn (16)*
	Johnson	*Scheuplein*	
Progesterone	−4.92	−6.38	−4.94
Testosterone	−6.21	−6.95	−5.67
Corticosterone	−7.08	−7.78	−6.88
Aldosterone	−7.79	−9.08	−7.86
Estradiol	−5.95	−7.08	−6.26
Hydrocortisone	−7.67	−9.08	−8.11
Pregnenolone		−6.38	−5.51
Hydroxyprogesterone-17a		−6.78	−5.73
Hydroxypregnenolone-17a		−6.78	−6.29
Deoxycorticosterone		−6.90	−5.59
Cortexolone		−7.68	−6.70
Cortisone		−8.56	−7.61
Estrone		−6.00	−5.66
Estratriol		−7.95	−7.78

also be estimated by the addition of fragment values (substructures). There remain three descriptors that have to be obtained, viz π_2^H, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$.

Firstly, we should point out that for a number of solutes, specifically some anilines, pyridines and sulfoxides (but not sulfones), the hydrogen-bond basicity appears to alter with the water–solvent system, as first observed by Leahy *et al*.[45] We have dealt with this problem by assigning two $\Sigma\beta_2$ values to these particular compounds, the original $\Sigma\beta_2^H$ value for use in systems where the solvent is very immiscible with water (eg alkanes, chloroform, benzene, etc) and an additional $\Sigma\beta_2^0$ descriptor for use in solvents that are somewhat miscible with water (eg octanol, ether, ethyl acetate). We note that this alternative descriptor is only needed for the specific solutes in the more miscible solvents, and that for partition between water and very immiscible solvents, the $\Sigma\beta_2^H$ descriptor is still used for the specific solutes. In the examples that follow, the solutes are all 'normal' solutes. There are a number of ways in which we have obtained the three required descriptors:

By analogy

Descriptors can be estimated by analogy with other compounds, especially in an homologous series. Except for the first one or two members, π_2^H, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ are invariably constant along an homologous series. In many other cases, one or more descriptors can be assigned by analogy.

By experiment

The best experimental method for the determination of descriptors is from a number of water–solvent log P values. If the coefficients in equation (1) are known for various water–solvent systems, then values of π_2^H, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ can be chosen so as to best reproduce the log P values. The choice of the partition systems is crucial; the water–solvent systems should have coefficients in equation (1) as different as possible. The coefficients in equation (1) are given in Table 8 for a number of water–solvent partitions. A suitable minimum set of three solvents is octanol, an alkane or cyclohexane, and a chlorinated hydrocarbon such as dichloromethane, trichloromethane or 1,2-dichloroethane. Octanol has a moderate *s*-coefficient, a zero *a*-coefficient and a large *b*-coefficient; alkanes have large *s*-, *a*- and *b*-coefficients, and the three above chlorinated solvents have low or zero *s*-coefficients and reasonably large *a*- and *b*-coefficients. However, if three solvent systems are chosen that have rather close *s*-, *a*- and *b*-coefficients, it will not be possible to obtain any reliable values of π_2^H, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$.

An example where a rather large number of log P values are available for a given solute is 4-aminobenzoic acid. In Table 9 are listed the log P values for the neutral species that we have used in our analysis. We have estimated R_2 as 1.075 by addition of fragment values, and have calculated V_x as 1.0315 in $(dm^3\,mol^{-1})/100$. The three remaining descriptors are then chosen to give the best fit between the observed log P values, and log P values calculated from the assigned descriptors and the coefficients in Table 8. As can be seen from Table 9, in this particular case it is possible to assign descriptors that will yield a set of calculated log P values that fit the observed log P values with an *sd* of 0.15 log units. The log P value for 4-aminobenzoic acid in the water–ethyl acetate system was out of line by 0.63 log units and was not used in the calculation. In our method of analysis, it is not at all unusual to find log P values out of line, and we think one useful advantage is the ability to identify out-of-line log P values.

Log P values in a large number of water–solvent systems are not always available, and descriptors for

Table 8. Coefficients in equation (1) for water–solvent partitions

Solvent	*c*	*r*	*s*	*a*	*b*	*v*
Isobutanol	0.249	0.480	−0.639	−0.050	−2.284	2.758
Pentanol	0.175	0.575	−0.787	0.020	−2.837	3.249
Octanol	0.088	0.562	−1.054	0.034	−3.460	3.814
Hexadecane	0.087	0.667	−1.617	−3.587	−4.869	4.433
Alkane	0.287	0.649	−1.657	−3.516	−4.818	4.282
Cyclohexane	0.127	0.816	−1.731	−3.778	−4.905	4.646
CCl_4	0.212	0.602	−1.234	−3.515	−4.528	4.552
$CHCl_3$	0.205	0.194	−0.412	−3.319	−3.455	4.403
CH_2Cl_2	0.314	0.001	0.022	−3.238	−4.137	4.259
CH_2ClCH_2Cl	0.227	0.278	−0.167	−2.816	−4.324	4.205
Benzene	0.017	0.490	−0.604	−3.013	−4.628	4.587
Toluene	0.015	0.594	−0.781	−2.918	−4.571	4.533
Et_2O	0.256	0.649	−1.130	−0.103	−4.998	4.380
n-Bu_2O	0.184	0.817	−1.495	−0.830	−5.090	4.694
EtOAc	0.253	1.157	−1.397	−0.054	−3.755	3.726
n-BuOAc	−0.468	0.712	−0.397	0.010	−3.743	3.865

Table 9. Calculation of descriptors for the neutral form of 4-aminobenzoic acid; R_2 taken as 1.075 and V_x calculated as 1.0315: best fit descriptors: $\pi_2^H = 1.57$, $\Sigma\alpha_2^H = 0.90$, $\Sigma\beta_2^H = 0.65$

Solvent	*Log P (calc)*	*Log P (obs)*
Isobutanol	1.077	0.89
Pentanol	1.083	0.90
Octanol	0.753	0.83
Alkane	−3.491	−3.74
Cyclohexane	−3.510	−3.25
CCl_4	−2.490	−2.48
$CHCl_3$	−0.973	−0.92
CH_2Cl_2	−0.860	− 0.80
Benzene	− 1.393	− 1.46
Et_2O	0.356	0.54
n-Bu_2O	− 0.499	− 0.43
n-BuOAc	1.237	1.17
EtOAc[a]	0.658	1.29

[a] Not used in the analysis.

the neutral forms of ampholytes have been obtained from just three log P values in the key solvents octanol, cyclohexane and dichloromethane.[46] In such cases, the determination of partition coefficients for the neutral species requires considerable extra information on macro- and micro-protonation constants in water. Other parameters can be used to calculate descriptors. For example, physicochemical measurements such as HPLC capacity factors can be used in addition to, or in place of, one or more water–solvent systems.

By fragment addition

In the absence of any measured log P or other physicochemical properties, it is possible to estimate descriptors by the addition of values for fragments (substructures) for which descriptors are available.[9,23] Indeed, for the prediction of descriptors, and hence the prediction of properties, from structure this is the only method that can be used.

4 CONCLUSIONS

It is possible to devise a set of physicochemical properties of solutes that correspond to given solute–solvent interactions. In particular, the solute properties include hydrogen-bond acidity and hydrogen-bond basicity, which are not the same as proton acidity and proton basicity, and so form new scales of acidity and basicity. The solute properties so constructed can be used as solute descriptors in a general equation for the correlation and prediction of transport-related properties that are important in areas as diverse as physical chemistry, pharmaceutical and medicinal chemistry, and environmental science.

Not only can the equations for transport-related properties be used to predict further values, but, unlike most correlative equations, those we have developed yield valuable information about the nature of the investigated processes. Thus it is now possible to deduce, for example, that the water–octanol system will not be a good model for blood–brain distribution, and to know exactly why it is not a good model.

As more and more processes are examined through our general equation, we should be in a position to set out a quite general method for the characterisation of these processes through the regression equation coefficients. This is an area in which we are actively working, in addition to extending the number and variety of solutes for which we have descriptors.

ACKNOWLEDGEMENTS

We thank Dr Mark Johnson for his helpful comments on the permeation of steroids, and the editor of the Journal of Pharmacy and Pharmacology for permission to reproduce Figures 5–8.

REFERENCES

1 Kamlet MJ, Doherty RM, Abboud J-LM, Abraham MH and Taft RW, Solubility: a new look. *Chemtech* **16**:566–576 (1986).
2 Abraham MH, Doherty RM, Kamlet MJ and Taft RW, A new look at acids and bases. *Chem Brit* **22**:551–554 (1986).
3 Abraham MH, Whiting GS, Doherty RM and Shuely WJ, Hydrogen bonding. Part 13. A new method for the characterisation of GLC stationary phases – the Laffort data set, *J Chem Soc Perkin Trans 2* 1451–1460 (1986).
4 Abraham MH, Whiting GS, Doherty RM and Shuely WJ, Hydrogen bonding. Part 14. The characterisation of some *N*-substituted amides as solvents: comparison with gas-liquid chromatography stationary phases. *J Chem Soc Perkin Trans 2* 1851–1857 (1986).
5 Abraham MH, Whiting GS, Donerty RM and Shuely WJ, Hydrogen bonding. Part 16. A new solute solvation parameter, π_2^H, from gas chromatographic data. *J Chromatogr* **587**:213–228 (1991).
6 Abraham MH, Hydrogen bonding. Part 31. Construction of a scale of solute effective or summation hydrogen-bond basicity. *J Phys Org Chem* **6**:660–684 (1993).
7 Abraham MH, Application of solvation equations to chemical and biochemical processes. *Pure App Chem* **65**:2503–2512 (1993).
8 Abraham MH, Scales of hydrogen-bonding: their construction and application to physicochemical and biochemical processes, *Chem Soc Revs* **22**:73–83 (1993).
9 Abraham MH and Chadha HS, Applications of a solvation equation to drug transport properties, in *Lipophilicity in Drug Action and Toxicology*, Ed. by Pliska V, Testa B and van de Waterbeemd H, VCH, Weinheim, Germany, pp 311–337 (1996).
10 Abraham MH and McGowan JC, The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. *Chromatographia* **23**:243–246 (1987).
11 Abraham MH, Andonian-Haftvan J, Whiting GS, Leo A and Taft RW, Hydrogen bonding. 34. The factors that influence the solubility of gases and vapours in water at 298 K, and a new method for its determination. *J Chem Soc Perkin Trans 2* 1777–1791 (1994).
12 Abraham MH, Chadha HS and Mitchell RC, Hydrogen

bonding. 32. An analysis of water–octanol and water–alkane partitioning and the Δlog P parameter of Seiler. *J Pharm Sci* 83:1085–1100 (1994).
13 Seiler P, Interconversion of lipophilicities from hydrocarbon/water systems into the octanol/water system. *Eur J Med Chem* **9**:473–479 (1974).
14 Abraham MH and Rafols C, Factors that influence tadpole narcosis. An LFER analysis. *J Chem Soc Perkin Trans 2* 1843–1851 (1995).
15 Overton CE, *Studien uber die Narkose*, Fischer, Jena (1901).
16 Overton CE, *Studies on Narcosis*, Ed by Lipnick RL, Chapman and Hall, London (1991).
17 Bradbury MWB, *The Concept of a Blood–Brain Barrier*, Wiley, New York (1979).
18 Hansch C, Steward AR, Anderson SM and Bentley DL, The parabolic dependence of drug action upon lipophilic character as revealed by a study of hypnosis. *J Med Chem* **11**:1–11 (1968).
19 Laduron PM, Janssen PFM, Gommeren W and Leysen JE, In-vitro and in-vivo binding characteristics of a new long-acting histamine H-1 antagonist, astemizole. *Mol Pharmacol* **21**:294–300 (1982).
20 Awouters FHL, Niemegeers CJE and Janssen PAJ, Pharmacology of the specific H_1 antagonist, astemizole. *Arzneim Forsch* **33**:381–388 (1983).
21 Young RC, Mitchell RC, Brown TH, Ganellin CR, Griffiths R, Jones M, Rana KK, Saunders D, Smith IR, Sore NE and Wilks TJ, Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H_2 receptor histamine antagonists. *J Med Chem* **31**:656–671 (1988).
22 Abraham MH and Weathersby PK, Hydrogen bonding. 30. Solubility of gases and vapours in biological liquids and tissues. *J Pharm Sci* **83**:1450–1456 (1994).
23 Abraham MH, Chadha HS and Mitchell RC, Hydrogen bonding. 33. Factors that influence the distribution of solutes between blood and brain. *J Pharm Sci* **83**:1257–1268 (1994).
24 Chadha HS, Abraham MH and Mitchell RC, Physicochemical analysis of the factors governing distribution of solutes between blood and brain. *Bioorg Med Chem Letters* **4**:2511–2516 (1994).
25 Abraham MH, Chadha HS and Mitchell RC, Hydrogen bonding. Part 36. Determination of blood brain distribution using octanol–water partition coefficients. *Drug Design Discovery* **13**:123–131 (1995).
26 van de Waterbeemd H and Kansy M, Hydrogen-bonding capacity and brain penetration. *Chimia* **46**:299–303 (1992).
27 Calder JAD and Ganellin CR, Predicting the brain-penetrating capability of histerminergic compounds. *Drug Design Discovery* **11**:259–268 (1994).
28 Salminen T, Pulli A and Taskinen J, Relationship between immobilised artificial membrane chromatographic retention and the brain penetration of structurally diverse drugs. *J Pharm Biomed Analysis* **15**:469–477 (1997).
29 Lombardo F, Blake JF and Curatoio WJ, Computation of brain–blood partitioning of organic solutes via free energy calculations. *J Med Chem* **39**:4750–4755 (1996).
30 Gratton JA, Lightman SL and Bradbury MW, Transport into retina measured by short vascular perfusion in the rat. *J. Physiol* **470**:651–663 (1993).
31 Gratton JA, Abraham MH, Bradbury MW and Chadha HS, Molecular factors influencing drug transfer across the blood–brain barrier. *J Pharm Pharmacol* **49**:1211–1216 (1997).
32 Timmermans PBMWM, Brands A and van Zwieten PA, Lipophilicity and brain disposition of clonidine and structurally related imidazolidines, *Arch Pharmacol* **300**:217–226 (1977).
33 Rapoport SI, Ohno K and Pettigrew KD, Drug entry into the brain. *Brain Res* **172**:354–359 (1979).
34 Pardridge WM, Triguero D, Yang J and Cancilla PA, Comparison of in-vitro and in-vivo models of drug transcytosis through the blood–brain barrier, *J Pharmacol Exp Therapeutics* **253**:884–891 (1990).
35 Brown SL and Rossi JE, A simple method for estimating dermal absorption of chemicals in water. *Chemosphere* **19**:1989–2001 (1989).
36 Fiserova-Bergerova V, Pierce JT and Droz PO, Dermal absorption potential of industrial chemicals: criteria for skin notation. *Am J Ind Med* **17**:617–635 (1990).
37 McKone TE and Howd RA, Estimating dermal uptake of nonionic organic chemicals from water and soil: I. unified fugacity-based models for risk assessments. *Risk Anal* **12**:543–557 (1992).
38 Guy RH and Potts RO, Penetration of industrial chemicals across the skin: a predictive model. *Am J Ind Med* **23**:711–719 (1993).
39 Wilschut A, ten Berge WF, Robinson PJ and McKone TE, Estimating skin permeation. The validation of five mathematical models. *Chemosphere* **30**:1275–1296 (1995).
40 Scheuplein RJ, Blank IH, Brauner MD and MacFarlane DJ, Percutaneous absorption of steroids. *J Invest Dermatol* **52**:63–70 (1969).
41 Johnson MF, Blankschtein D and Langer R, Permeation of steroids through human skin. *J Pharm Sci* **84**:1144–1146 (1995).
42 Abraham MH, Chadha HS and Mitchell RC, The factors that influence skin penetration of solutes, *J Pharm Pharmacol* **47**:8–16 (1995).
43 Abraham MH, Martins F and Mitchell RC, Algorithms for skin permeability using hydrogen bond descriptors: the problem of steroids. *J Pharm Pharmacol* **49**:858–865 (1997).
44 Anderson BD, Higuchi WI and Raykar PV, Heterogeneity effects on permeability–partition coefficient relationships in human stratum corneum. *Pharm Res* **5**:566–573 (1988).
45 Leahy DE, Morris JJ, Taylor PJ and Wait AR, Model solvent systems for QSAR. Part 3. An LSER analysis of the 'critical quartet'. New light on hydrogen bond strength and directionality. *J Chem Soc Perkin Trans 2* 705–722 (1992).
46 Abraham MH, Takacs-Novak K and Mitchell RC, On the partition of ampholytes: application to blood–brain distribution. *J Pharm Sci* **86**:310–315 (1997).